WHAT IS CLAIMED IS:

1.    A method for recovery from failures affecting a primary copy of a data repository, for use in a data processing system in which updates applied to the data repository during normal forward processing are applied within transactional units of work, the method including the steps of:

     storing a secondary copy of data representing data items held within the data repository and updates applied to the data repository within said units of work;

     in response to a failure affecting a primary copy of the data repository, identifying from said secondary copy a set of operations required for restoring said data items and applied updates to a primary copy of the data repository;

     determining the state, at the time of the failure, of each unit of work corresponding to one or more operations of the identified set of restore operations; and

     performing restore operations of said identified set for which said performance is consistent with the determined state of the corresponding unit of work, and discarding restore operations of said identified set for which performance is inconsistent with the determined state of the corresponding unit of work.

2.    A method according to claim 1, including the steps of:

saving to a cache a subset of said secondary copy of
data, which subset corresponds to the identified set of
operations required for restoring said data items and
applied updates;

5          and wherein, subsequent to the step of determining
the state of each unit of work, the step of performing
restore operations comprises applying restore operations
from said cache.


10    3.   A method according to claim 2, including the step of
deleting from the cache the restore operations for which
the corresponding unit of work is determined to be
neither committed nor in-doubt, thereby to discard said
restore operations for which performance is inconsistent

15    with the determined state of the corresponding unit of
work, when performing restore operations.


4.   A method according to claim 1, wherein the step of
performing restore operations includes the steps of:

20          performing restore operations for which the
corresponding unit of work is determined to be committed;
and

          performing restore operations for which the
corresponding unit of work is determined to be in-doubt,

25    and marking the data item to indicate that the unit of
work is in-doubt.


5.   A method according to claim 2, including the step of
deleting from the cache any pairs of updates within the

30    set of restore operations, which pair of updates

correspond to addition of a data item and retrieval of
the same data item and which pair of updates was
completed prior to the failure, thereby to discard said
pairs of updates when performing restore operations.

6.    A method according to claim 1, wherein storing the
secondary copy comprises storing a backup copy of the
data repository and storing log records describing
updates to the primary copy performed since the backup
copy was stored; and wherein the step of identifying said
set of operations comprises replaying the log records to
identify operations performed on the primary copy of the
data repository.

7.    A method according to claim 1, wherein storing the
secondary data copy includes maintaining log records that
describe operations performed on data items within the
data repository, and wherein the step of restoring data
to the primary copy of the data repository includes the
steps of:
        replaying the log records of operations performed on
data items within the data repository,
        caching log records relating to operations performed
on data items within the data repository within an
original unit of work,
        determining from the cached log records the state of
the original units of work at the time of the failure,
and
        determining, for said operations having cached log
records, which operations to perform within the recovery

unit of work based on the determined state of the
original units of work.

8.    A method according to claim 1, wherein the data
repository is a message repository and the step of
restoring data to the primary copy of the data repository
comprises performing message add, update and delete
operations on the message repository.

9.    A method according to claim 8, for performance
within a messaging communication system, wherein
maintaining the secondary data copy includes storing log
records to describe updates to the primary copy, and
wherein the step of restoring data to the primary copy of
the repository includes the steps of caching log records
relating to message add, update and delete operations
performed under syncpoint control within an original unit
of work, determining from the log records the state of
the original unit of work at the time of the failure, and
determining the operations to perform within the recovery
unit of work based on the determined state of the
original unit of work as follows:
        if the original unit of work is committed,
performing the relevant message add, update and delete
operations; and
        if the original unit of work is in-doubt, performing
the relevant message add, update and delete operations
but marking the operations in-doubt; and
        if the original unit of work is neither committed
nor in-doubt, discarding the cached operations.

10.  A data communication system including:

data storage for storing a primary copy of a data
repository;

secondary data storage for storing a secondary copy
of data representing the data repository which secondary
data is sufficient to recover the primary copy of the
data repository and data held thereon;

` a recovery component for controlling the operation
of the data communication system to recover from a
failure affecting the primary copy of the data
repository, wherein the recovery component is operable to
control the data communication system  to perform the
steps of:

in response to a failure affecting a primary copy of
the data repository, identifying from said secondary copy
a set of operations required for restoring said data
items and applied updates to a primary copy of the data
repository;

determining the state, at the time of the failure,
of each unit of work corresponding to one or more
operations of the identified set of restore operations;
and

performing restore operations of said identified set
for which said performance is consistent with the
determined state of the corresponding unit of work, and
discarding restore operations of said identified set for
which performance is inconsistent with the determined
state of the corresponding unit of work.

11.   A data communication system for transferring
messages between a sender and a receiver, the system
including data storage for storing a primary copy of a
message repository and including secondary data storage,
wherein messages are held in the primary copy of the
message repository following a message send operation and
are retrieved from the primary copy of the message
repository for delivery to the receiver, and wherein a
secondary copy of the message repository is stored in the
secondary data storage and log records are written to
record message send and message retrieval events
performed within transactional units of work since
creation of the secondary copy,

the system including a recovery component adapted to
control the data communication system to perform the
following steps:

in response to a failure affecting a primary copy of
the message repository, identifying from said secondary
copy a set of operations required for restoring said
messages and reapplying message send and retrieval
operations to a primary copy of the message repository;

determining the state, at the time of the failure,
of each unit of work corresponding to one or more
operations of the identified set of restore operations;
and

performing restore operations of said identified set
for which said performance is consistent with the
determined state of the corresponding unit of work, and
discarding restore operations of said identified set for

which performance is inconsistent with the determined
state of the corresponding unit of work.


12.    A computer program product comprising program code
recorded on a recording medium for controlling the
operation of a data processing apparatus on which the
program code executes to perform a method for recovering
a data repository from a failure affecting a primary copy
of the data repository, for use with a data processing
apparatus having a secondary data storage and having a
component for maintaining a secondary copy of data in the
secondary data storage which secondary copy is sufficient
to recover the primary copy of the data respository and
data items held thereon, and wherein updates applied to
the data repository are applied within transactional
units of work, the method including the steps of:
      in response to a failure affecting a primary copy of
the data repository, identifying from said secondary copy
a set of operations required for restoring said data
items and applied updates to a primary copy of the data
repository;
      determining the state, at the time of the failure,
of each unit of work corresponding to one or more
operations of the identified set of restore operations;
and
      performing restore operations of said identified set
for which said performance is consistent with the
determined state of the corresponding unit of work, and
discarding restore operations of said identified set for

which performance is inconsistent with the determined
state of the corresponding unit of work.

13.   A recovery component for recovering a data
repository from a failure affecting a primary copy of the
data repository, for use with a data processing system
having primary and secondary data storage and having a
component for maintaining a secondary copy of data in the
secondary data storage which secondary copy is sufficient
to recover the primary copy of the data respository and
data items held thereon, wherein updates applied to the
data repository are applied within transactional units of
work, the recovery component being adapted to perform a
method including the steps of:
        in response to a failure affecting a primary copy of
the data repository, identifying from said secondary copy
a set of operations required for restoring said data
items and applied updates to a primary copy of the data
repository;
        determining the state, at the time of the failure,
of each unit of work corresponding to one or more
operations of the identified set of restore operations;
and
        performing restore operations of said identified set
for which said performance is consistent with the
determined state of the corresponding unit of work, and
discarding restore operations of said identified set for
which performance is inconsistent with the determined
state of the corresponding unit of work.